



Multidimensional Clusters of CD4+T Cell Dysfunction Are Primarily Associated with the CD4/CD8 Ratio in Chronic HIV Infection

Frederiksen, Juliet Wairimu; Buggert, Marcus; Noyan, Kajsa; Nowak, Piotr; Sönnernborg, Anders; Lund, Ole; Karlsson, Annika C.

Published in:
P L o S One

Link to article, DOI:
[10.1371/journal.pone.0137635](https://doi.org/10.1371/journal.pone.0137635)

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Frederiksen, J. W., Buggert, M., Noyan, K., Nowak, P., Sönnernborg, A., Lund, O., & Karlsson, A. C. (2015). Multidimensional Clusters of CD4+T Cell Dysfunction Are Primarily Associated with the CD4/CD8 Ratio in Chronic HIV Infection. *P L o S One*, 10(9), [e0137635]. <https://doi.org/10.1371/journal.pone.0137635>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

RESEARCH ARTICLE

Multidimensional Clusters of CD4+ T Cell Dysfunction Are Primarily Associated with the CD4/CD8 Ratio in Chronic HIV Infection

Juliet Frederiksen^{1☯}, Marcus Buggert^{2,3☯*}, Kajsa Noyan², Piotr Nowak^{2,4}, Anders Sönnnerborg^{2,4}, Ole Lund¹, Annika C. Karlsson²

1 Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark, **2** Division of Clinical Microbiology, Department of Laboratory Medicine, Karolinska Institutet, Stockholm, Sweden, **3** Department of Microbiology, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, **4** Unit of Infectious Diseases, Department of Medicine Huddinge, Karolinska Institutet, Karolinska University Hospital Huddinge, Stockholm, Sweden

☯ These authors contributed equally to this work.

* marcus.buggert@ki.se



OPEN ACCESS

Citation: Frederiksen J, Buggert M, Noyan K, Nowak P, Sönnnerborg A, Lund O, et al. (2015) Multidimensional Clusters of CD4+ T Cell Dysfunction Are Primarily Associated with the CD4/CD8 Ratio in Chronic HIV Infection. PLoS ONE 10(9): e0137635. doi:10.1371/journal.pone.0137635

Editor: William A Paxton, Institute of Infection and Global Health, UNITED KINGDOM

Received: May 26, 2015

Accepted: July 30, 2015

Published: September 24, 2015

Copyright: © 2015 Frederiksen et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by The Swedish Research Council (Grants K2010-56X-20345-04-3 and K2014-57X-22451-01-5), <http://www.vr.se/>; The Swedish Physicians Against AIDS Research Foundation (Fob2013-0008), <http://www.aidsfond.se/>; Magnus Bergvalls foundation (2013), <http://www.magnbergvallsstiftelse.nu/>; Karolinska Institutet (Partial funding of new postgraduate student at Karolinska Institutet 2010–2013), <https://internwebben.ki.se/en/grants-office>.

Abstract

HIV infection provokes a myriad of pathological effects on the immune system where many markers of CD4+ T cell dysfunction have been identified. However, most studies to date have focused on single/double measurements of immune dysfunction, while the identification of pathological CD4+ T cell clusters that is highly associated to a specific biomarker for HIV disease remain less studied. Here, multi-parametric flow cytometry was used to investigate immune activation, exhaustion, and senescence of diverse maturation phenotypes of CD4+ T cells. The traditional method of manual data analysis was compared to a multidimensional clustering tool, FLOW Clustering with K (FLOCK) in two cohorts of 47 untreated HIV-infected individuals and 21 age and sex matched healthy controls. In order to reduce the subjectivity of FLOCK, we developed an “artificial reference”, using 2% of all CD4+ gated T cells from each of the HIV-infected individuals. Principle component analyses demonstrated that using an artificial reference lead to a better separation of the HIV-infected individuals from the healthy controls as compared to using a single HIV-infected subject as a reference or analyzing data manually. Multiple correlation analyses between laboratory parameters and pathological CD4+ clusters revealed that the CD4/CD8 ratio was the pre-eminent surrogate marker of CD4+ T cells dysfunction using all three methods. Increased frequencies of an early-differentiated CD4+ T cell cluster with high CD38, HLA-DR and PD-1 expression were best correlated ($Rho = -0.80$, $P \text{ value} = 1.96 \times 10^{-11}$) with HIV disease progression as measured by the CD4/CD8 ratio. The novel approach described here can be used to identify cell clusters that distinguish healthy from HIV infected subjects and is biologically relevant for HIV disease progression. These results further emphasize that a simple measurement of the CD4/CD8 ratio is a useful biomarker for assessment of combined CD4+ T cell dysfunction in chronic HIV disease.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Human immunodeficiency virus type 1 (HIV) infection is characterized by an initial loss of CCR5+CD4+ T cells at mucosal sites of the body [1, 2], and later a gradual decline of central and effector memory CD4+ T cells due to high cell turnover [3], pyroptosis [4], apoptosis [5–7] and/or many other effects that impair normal immune homeostasis [3, 8–10]. Except from becoming infected with HIV, CD4+ T cells also exhibit numerous pathological changes that are contributors, or consequences, of HIV disease progression. The most classically studied markers of disease progression probably involve CD38 and HLA-DR, which are used as measurements of T cell activation [11]. Immune activation has previously been shown to be highly predictive of HIV disease progression [3, 12] and thought to be directly involved in the process of CD4+ T cell division and depletion [13, 14]. Importantly, CD38 and HLA-DR are elevated in most individuals on long-term combined antiretroviral therapy (ART) and predictive of immune recovery and mortality post ART [15]. PD-1 and Tim-3 are markers of T cell exhaustion where both have been shown to be elevated in dysfunctional T cells after HIV and other chronic viral infections [16–18]. Particularly elevated levels of PD-1, together with CD38 and HLA-DR expression, has previously been demonstrated in European [19] and African [20] cohorts to be highly associated with HIV disease progression, independently of T cell maturation phenotypes. Likewise, the memory phenotypes of CD4+ T cells might be highly skewed, where particularly markers of immunosenescence (CD28- and CD57+ cells) are upregulated in HIV-infected subjects, leading to poor T cell proliferation and homeostasis [21].

As HIV infection primarily affects CD4+ T cells, bulk measurements from heterogeneous samples with *e.g.* microarray prevent detailed characterization of immunological sub-populations [22]. Therefore, single-cell analysis tools, such as flow cytometry (FCM), are optimal for cell characterization within HIV research. Advances in the instrumentation and reagents have allowed an increased number of parameters to be measured simultaneously on individual cells, yielding data of high dimensionality and complexity [23–25]. Data analysis of high-dimensional FCM data has long been the caveat of polychromatic flow cytometry experiments [26, 27], where traditional data analysis techniques are not only time-consuming but also highly subjective to the experience of the operator [27–29].

Over the last decade though, there has been an increase in the number of automated data analysis techniques developed for FCM [30, 31]. The approaches can be divided into algorithms for automated subset identification and algorithms that quantify differences between multivariate distributions [32, 33]. The automated subset identification solutions are model-based, graph-based or multidimensional clustering approaches [30, 31]. The model-based methods use applications of Bayes mixture Gaussian models, t-mixture models and skew-t models [30]. These methods show robustness that follows predefined models, but they might be slow at estimating model parameters of high-dimensional dataset. SamSPECTRAL [34] is an example of the graph-based methods, while multidimensional clustering approaches available include SPADE [27], flowMeans [35] and FLOCK [36]. Here, we investigated whether one of the existing multidimensional clustering approaches, FLOCK, could delineate HIV-infected *versus* healthy control subjects based on the measurement of the eight markers of CD4+ T cell memory (CD45RO, CD27), activation (CD38, HLA-DR), exhaustion (PD-1, Tim-3) and senescence (CD28, CD57). FLOCK was chosen for these analyses due to its ability to determine centroids of a reference sample and apply these to other samples, and therefore making the results directly comparable between all samples. After developing an “artificial reference” patient, using gated CD4+ T cell data from the HIV-infected subjects; we investigated if FLOCK was 1) superior in delineating the HIV-infected from the healthy control

group compared to normal manual data analysis and 2) which traditional HIV disease bio-marker that was associated with the multidimensional clusters of CD4+ T cells dysfunction in chronic HIV disease.

Materials and Methods

Ethical statement

The Regional Ethical Review Board (Stockholm, Sweden, Dnr 2009-1485-31-3) approved the study. Written informed consent of all study subjects was documented in accordance with the Declaration of Helsinki and all participants were provided with written and oral information about the study.

Study subjects

In total, 47 HIV-infected individuals (HIV+) were recruited from the Karolinska University Hospital Huddinge, Stockholm, Sweden. All patient samples were collected from untreated subjects; except for three individuals with AIDS defining illnesses (AIDS patients), which viral load measurements were excluded from the statistical analysis. An age and sex matched healthy control group (n = 21) was recruited to compare the CD4+ T cell clusters (Table 1). This cohort has partly been used in a previous study by Buggert et al [19], where detailed information about the HIV-infected subjects are described [19].

Antibody reagents

A single flow cytometry panel was tested on all HIV-infected individuals and healthy control subjects as seen in Buggert et al. The antibodies used in the assay were: anti-CD3 APC-H7 (Clone SK7), anti-CD14 V500 (Clone M5E2), anti-CD19 V500 (Clone B43) (BD Bioscience); anti-CD45RO ECD (clone UCHL1) (Beckman Coulter); anti-CD4 BV650 (Clone OKT4), anti-PD-1 BV421 (clone EH12.2H7), anti-CD27 BV785 (clone O323), anti-CD28 PerCP-Cy5.5 (clone CD28.2), anti-CD38 APC (clone HIT2), anti-CD57 FITC (clone HCD57) (Biolegend); anti-HLA-DR PE-Cy7 (clone LN3) (eBioscience); anti-CD8 Qd565 (Clone 3B5) (Life Technologies) and anti-Tim-3 PE (clone 344823) (R&D Systems). LIVE/DEAD Aqua amine dye (Life Technologies) was used to discriminate dead cells and debris.

Table 1. Summary of the HIV-infected cohort.

Character	HIV-infected cohort
n	47
Age, years	40 (33–49)
Males, n (%)	31 Males (66%)
Females, n (%)	16 Females (33%)
CD4 counts, cells/ μ l	390 (310–500)
CD4%	23 (17–31)
CD8 counts, cells/ μ l	900 (610–1211)
CD8%	49 (43–60)
CD4/CD8 ratio	0.44 (0.27–0.77)
Viral load, copies/ml	33800 (1138–990000)

Median (IQR) shown for all parameters; n: numbers.

doi:10.1371/journal.pone.0137635.t001

Cells and flow cytometric staining

Peripheral blood mononuclear cells (PBMC) were isolated from EDTA collected whole blood by Hypaque-Ficoll (GE Healthcare) density gradient centrifugation and then cryopreserved in 90% FBS (Life Technologies) supplemented with DMSO (Sigma Aldrich). The cells were thawed, washed and rested overnight at 37°C, in R10 (RPMI-1640 Medium AQmedia (Sigma Aldrich) containing 10% FBS, 50 IU/mL penicillin and 50 µg/mL streptomycin, and 10 mM HEPES (Life Technologies)). The PBMCs were counted the next day using a Nucleocounter (ChemoMetec A/S) and added to a concentration of 1.5×10^6 cells/well in V-bottom plates. Cells were washed in PBS, containing 2mM EDTA, and stained with all antibody reagents for 30 min at 20°C as previously described [37]. Cells were washed and resuspended in PBS containing 1% paraformaldehyde (PFA). All flow cytometric analyses were conducted within 2hrs after fixation.

Flow cytometric analyses

PBMC were analysed on a LSR Fortessa (BD Biosciences) where minimally 600,000 total events were collected per run. Data compensation was performed using antibody capture beads (BD Biosciences) after separate stainings with all antibodies used in the experiment. FlowJo 8.8.7 (Treestar) were used for flow cytometric gating analyses and all manual gates were based on fluorescence minus one (FMO) staining [38].

Generation of the artificial reference

The manual gating of the FCM data produced CD4+ T cell gated data. This data was used to create the artificial patient. The median number of CD4+ T cells was calculated from the HIV+ data files. The desirable percentage events that were needed to be collected from each of the HIV+ data files was calculated, $\frac{\text{median}(\#CD4+ \text{ events})}{\text{total}(\#CD4+ \text{ events})} * 100$. Thereby, a random 2% of the files were collected to create an artificial HIV+ subject. The artificial HIV+ subject was used to create the centroid table and this was applied to the entire dataset, including both the HIV+ and control data files.

Data analysis

Heat maps in conjunction with unsupervised hierarchical clustering were used as a method for data visualization. The distance measure used for the hierarchical clustering is the dissimilarity index $dist_{i,j} = 1 - \text{cor}(p_i, p_j)$, where cor is the Spearman rank correlation calculated for p_i and p_j is the population frequency vector for subject i and j , respectively. The Kolmogorov-Smirnov (KS) test statistic is used to quantify the differences between two sample distributions to determine whether they were drawn from the same distribution. Experimental variables between healthy controls and HIV-infected individuals were analyzed using Mann-Whitney U test. Correlations were assessed using non-parametric Spearman rank tests. Bonferroni corrections were applied to all cases where multiple testing was performed. To summarize, flow cytometry data was analysed with FlowJo 8.8.7 (Treestar), FLOCK (ImmPort, NIH web site www.immport.org) and R environment [39].

Results

FLOCK was used for automated population (cluster) identification to analyze an eight-parameter dataset of HIV+ and healthy control subjects. The eight measured parameters were selected to give an overview of the activation (CD38, HLA-DR), exhaustion (PD-1, Tim-3), senescence (CD28, CD57) and memory differentiation (CD45RO, CD27) status of CD4+ T cells. A

traditional gating strategy for isolating CD4+ T cells were performed on the multiparametric flow cytometry data set prior to FLOCK data examination (Fig 1A). Thereafter, FLOCK analyses were performed on all of the HIV-infected subjects, where the number of clusters automatically identified by FLOCK differed between the subjects, ranging from 12–23 clusters. In the interest of being able to compare the data, a reference subject that identified the biologically relevant cell clusters was chosen and the centroids of these clusters were then applied to the remaining healthy controls and HIV-infected subjects. The unique clusters identified in the eight-dimensional space for the representative subject is shown in Fig 1B.

The development of an artificial reference to reduce the subjectivity

The need for a reference subject in the multidimensional clustering analyses is a source of bias in the FLOCK algorithm, as there is no automated method of selecting the optimal reference subject. Instead as a first step using FLOCK, a single subject is selected to function as a reference in the FLOCK analysis. To remove this subjectivity, a “mosaic” subject was created as an artificial control from a subset of HIV-infected subjects in the study. A random 2% of events from each of the 47 HIV-infected individuals were extracted and concatenated to create the artificial patient. This was hypothesized to give a representative view of the relative phenotypes found in the HIV-infected subjects. The following abbreviations will be used for the remainder of the article; the FLOCK analysis using the specific reference subject that identified the biologically relevant cell clusters will be referred to as sFLOCK and the FLOCK analysis using the artificial reference will be referred to as aFLOCK.

The distinctive clusters identified by sFLOCK (Fig 1B) yielded biologically accountable clusters, similar to those observed in aFLOCK (Fig 1C). The number of bins automatically determined by FLOCK for aFLOCK was greater than that found for sFLOCK, which leads to a larger number of aFLOCK clusters ($n = 25$) compared to sFLOCK ($n = 21$) as well as the largest number of clusters determined by any individual ($n = 23$). This agrees with our expectations, as the method would be sampling from a larger repertoire of cells. The clusters obtained from sFLOCK were shown to be tighter, where the synonymous clusters were more diffuse, which is most likely due to the events being an artificial drawn from different samples. Additionally, the aFLOCK clusters were shown to make biological sense and not artifacts from combining the different samples.

aFLOCK provides a better separation of HIV-infected subjects and healthy controls

In order to assess the performance of aFLOCK, it was compared to the results of the sFLOCK and manual data analysis. The format of the manual data analysis results is a frequency table with rows of the subjects used in the study and columns for the cell populations investigated (e.g. CD4+CD38+HLA-DR+, CD4+CD57+CD28-, etc). The results of the FLOCK analyses are tables of cell population frequencies for all study subjects.

Unsupervised hierarchical clustering in conjunction with heat maps, a method for the visualization of numeric matrices to find patterns in data in an unbiased fashion, was used to analyze the three data matrices containing the cluster frequency for row (individual) i and column (cluster) j . A visualization of the results of the manual and FLOCK FCM data analysis can be seen in Fig 2A–2C. As can visually be noted, the hierarchical clustering of the manual data analysis results failed to separate the HIV-infected individuals and the healthy controls (Fig 2A) as good as sFLOCK (Fig 2B) and aFLOCK (Fig 2C). Notably, the AIDS patients, and an additional HIV-infected subject, were distinguished as outliers in all three methods. It was the

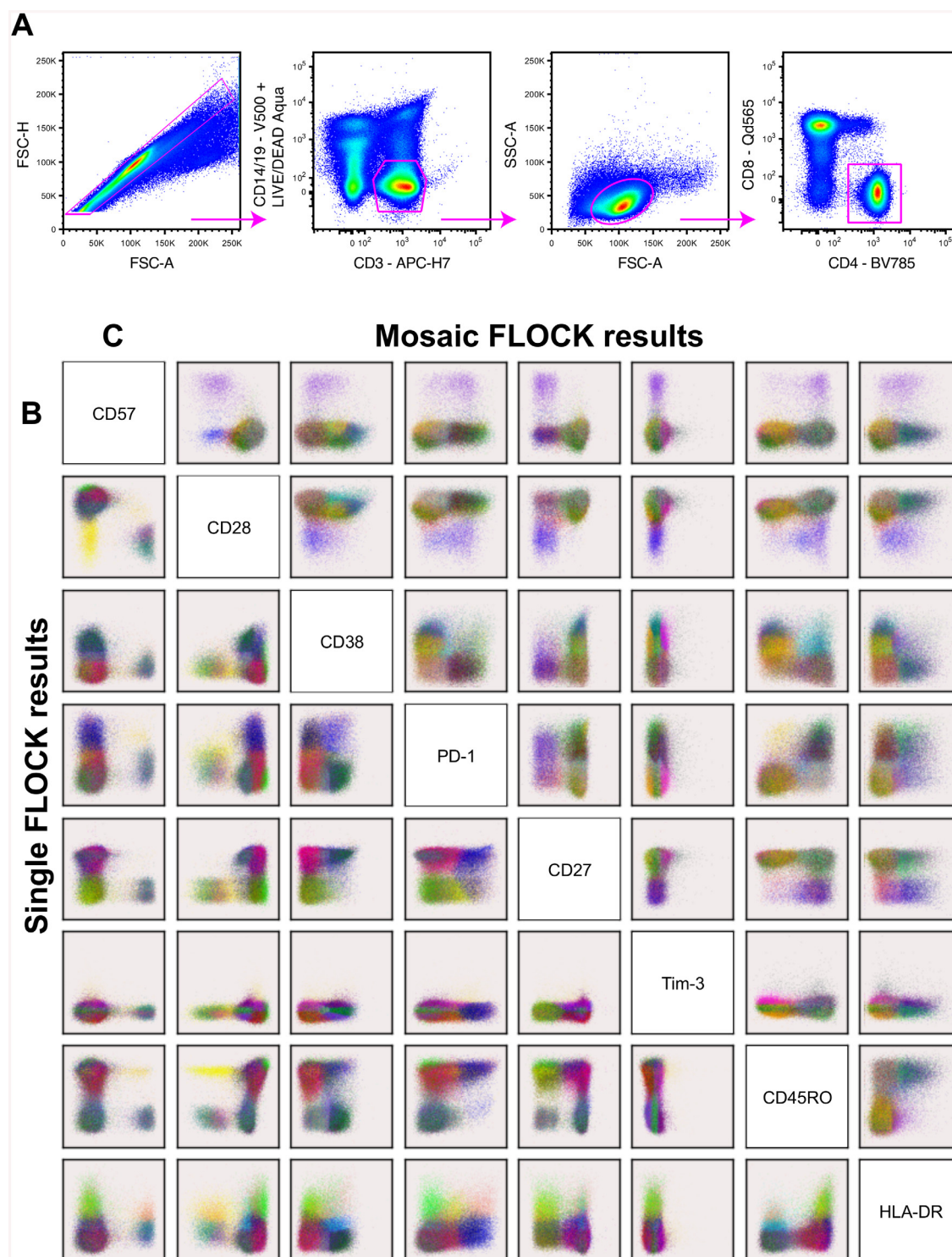


Fig 1. Flow cytometry gating and FLOCK populations. The manual gating strategy used to gate for the CD4+ T cells is shown in the top panel (A). The CD4+ T cell events were uploaded to immPort (import.niaid.nih.gov) for FLOCK analysis. The unique populations identified by FLOCK using the single HIV reference (bottom left) (B) and artificial reference (top right) (C) is shown. The artificial reference is made-up of 2% of the CD4+ T cells from each subject in the HIV cohort, whereas the single HIV reference is made-up of the CD4+ T cells from a single individual from the HIV cohort that appeared to be biologically representative of the cohort.

doi:10.1371/journal.pone.0137635.g001

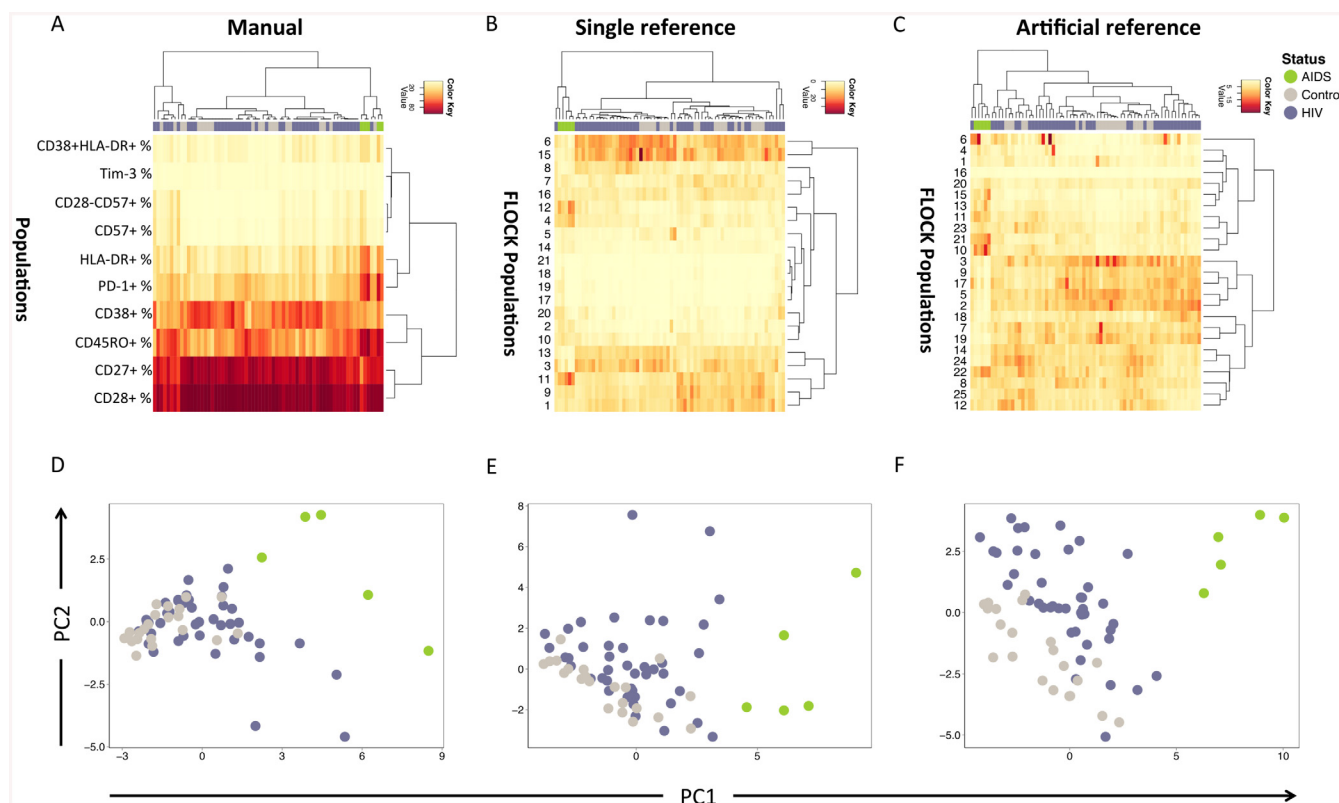


Fig 2. Clustering of HIV-infected and -uninfected subjects with manual and FLOCK gating principles. The top panel shows the heat map representation of the matrices containing the cell population frequencies of the manual gating results (A), the FLOCK results using one HIV infected subject that identified biologically relevant cell populations (B) and the FLOCK results using an artificial of the HIV-infected subjects as a reference (C). The bottom panel shows the principle component analysis (PCA) was performed on the matrices illustrated in A-C to investigate whether there were difference between the control, HIV-infected and AIDS subjects. The results of the PCA performed on the manually determined population frequencies is shown in (D), the results of the K-S test that compared the HIV infected individuals to the healthy controls for PC1 (P value = 0.0009, D value = 0.495) and PC2 (P value = 0.3, D value = 0.236) are shown below the biplot. The FLOCK data using one HIV infected subject that identified biologically relevant cell populations is shown in (E), the results of the K-S test that compared the HIV infected individuals to the healthy controls for PC1 (P value = 0.04, D = 0.353) and PC2 (P value = 0.02, D value = 0.378) are shown below the biplot. The FLOCK data using an artificial of the HIV-infected subjects as a reference is shown in (F), the results of the K-S test that compared the HIV infected individuals to the healthy controls for PC1 (P value = 0.02, D value = 0.384) and PC2 (P value = 0.0008, D value = 0.497) are shown below the biplot. A detailed overview of the FLOCK populations in (B) and (C) can be seen in [S1 Table](#).

doi:10.1371/journal.pone.0137635.g002

same HIV-infected subject, with a low CD4 count and CD4/CD8 ratio comparable to some of the AIDS patients that clustered with the AIDS patients in the FLOCK analyses.

The majority of healthy controls in the aFLOCK results ([Fig 2C](#)) were found to cluster in two groups, whereas the sFLOCK ([Fig 2B](#)) and manual data analysis ([Fig 2A](#)) results showed the clustering of healthy controls to be more spread in comparison. The HIV-infected individual found to cluster with the AIDS patients in the manual data analysis ([Fig 2A](#)) had a relatively high CD4/CD8 ratio. In addition to these findings, a healthy control was surprisingly found within this AIDS cluster for the manual data analysis. Overall, these results thus suggest that the aFLOCK captures an immunopathological signal that can more easily separate the HIV-infected individuals from the healthy controls.

To objectively compare the results from the three different FCM data analysis methods, principle component analysis (PCA) of the frequency tables were performed ([Fig 2D–2F](#)). It was clearly observed that the AIDS patients were outliers in the PCA biplots of both FLOCK data analyses methods ([Fig 2E and 2F](#)), but particularly remarkable in the aFLOCK analysis ([Fig 2F](#)). Kolmogorov-Smirnov (KS) tests were used to quantify the distance between two

groups to determine whether they were drawn from the same distribution (P-value of 1 for this case). The KS test showed that the distribution of PC1 scores between HIV-infected individuals and control subjects (aFLOCK analysis, [Fig 2F](#)) were significantly different (P value = 0.02, D value = 0.384). Similarly for PC2 and the results of the KS test revealed that the HIV+ and control subjects' distributions were highly significantly different (P value = 0.0008, D value = 0.497). More specifically, it was clusters 5, 10, 2, 11 and 21 ([S1 Table](#)) that had the greatest impact on the PC1 scores. These clusters were split between naïve T cell clusters (clusters 2 and 5) and exhausted (\pm activated) memory T cell clusters (clusters 10, 11 and 21). The PC2 dimension was more efficient at separating the HIV-infected and control subjects, where the clusters that had the largest impact on PC2 were 24, 14, 25, 20 and 12 (mostly naïve T cell clusters). The top 5 clusters that explain PC2 were seen to adhere together in the right part of the heat map, whereas those for PC1 were seen to cluster together in the middle. The KS test for the PCA results of sFLOCK ([Fig 2E](#)) showed a significant difference (P value = 0.04, D = 0.353) for the PC1 scores when comparing the distributions of the HIV+ subjects and the healthy controls. A significant difference (P value = 0.02, D value = 0.378) was also seen for the KS test of PC2. In the case of the PCA of the manual FCM data analysis, the KS-test of the PC1 scores for the two groups showed a significant difference (P value = 0.0009, D value = 0.495), however, the PC2 distributions for the two groups did not show a significant difference (P value = 0.3, D value = 0.236).

The CD4/CD8 ratio is primarily associated with the immunopathogenic clusters

The FLOCK and manually gated clusters were next correlated to numerous routine laboratory parameters that are used to assess the immune status of HIV-infected individuals. The initial step of the analysis involved the selection for the immunopathogenic clusters for each data set. This was done by performing multiple Mann-Whitney U tests to select for clusters that varied significantly, after Bonferroni correction for multiple testing. The pre-processing step resulted in six clusters (ten clusters before Bonferroni adjustments) for the manual dataset, seven clusters (21 clusters before Bonferroni adjustments) for the sFLOCK dataset and twelve clusters (25 clusters before Bonferroni adjustments) for the aFLOCK dataset. The phenotypes of the sFLOCK and aFLOCK results are shown in [S1 Table](#). Notably, almost half of the original artificial reference dataset captured an immunopathogenic difference between the HIV+ and healthy control groups, whereas only a third of the sFLOCK dataset captured this immunopathogenesis. Box plot representations of these multiple Mann-Whitney U tests are illustrated in [Fig 3](#).

Multiple non-parametric Spearman correlation tests between the immunopathogenic clusters and the routine laboratory parameters were next employed ([Table 2](#)). The detailed overview of these results is observed in [S2](#) and [S3](#) Tables. It was observable that the rank of the average absolute Rho-values for the different methods corresponds very well to each other. The average absolute Rho-values were calculated using the union of the immunopathological clusters significantly correlated to any of the clinical markers. In each of the three methods, the CD4/CD8 ratio were shown to have the greatest correlation to the manual or FLOCK gated clusters ([Table 2](#)). A close second in each dataset was the CD4%. The CD4% and CD4/CD8 ratio had a correlation coefficient of 0.98 (P value = 2.2×10^{-16}) due to the fact that the formula to calculate them differs by an additional denominator, explaining the reason why the average spearman correlation coefficients were very similar.

The T cell clusters 15, 21, 10, 11 and 13 significantly correlated with at least one of the laboratory parameters in the aFLOCK dataset. These clusters contained cells that were CD28+,

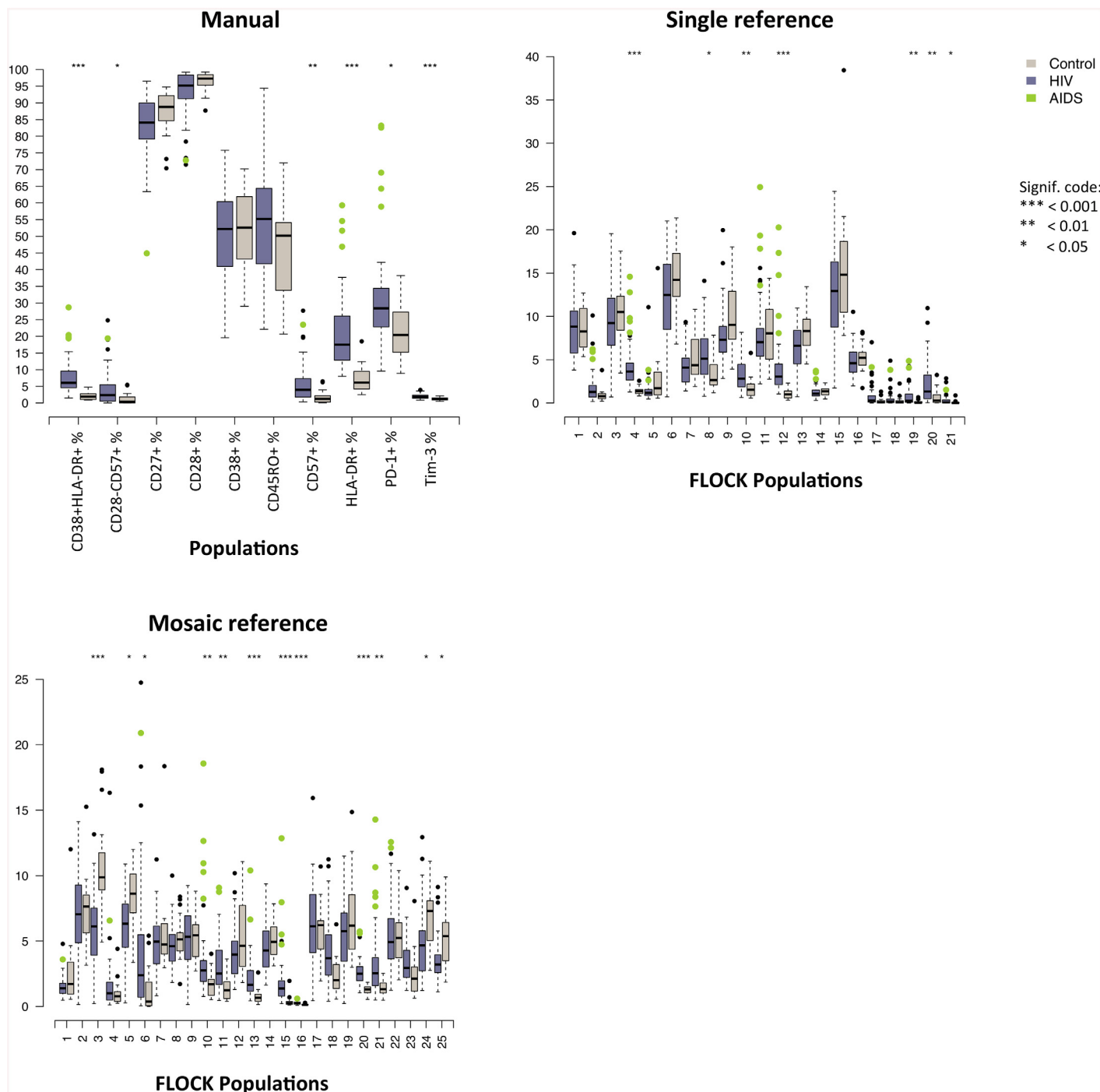


Fig 3. Box plots of all manual and FLOCK gated populations between HIV-infected and -uninfected subjects. Box plot representation of summary results of the two groups generated by the three methods used to investigate the same HIV immunopathogenesis dataset, manual data analysis (left), FLOCK data analysis using the single HIV reference (middle) and FLOCK data analysis using the artificial reference (right). The data is presented in a box and whisker plot where the horizontal line in the box is the median population occupation, the edges of the boxes are the 25th and 75th percentiles of the population occupation and the 'whiskers' represent the 10th and 90th percentiles of the population occupation, and the dots indicate outliers. The purple and grey boxes represent the HIV+ and healthy control group, respectively, where the green dots indicate outliers that are AIDS patients. Results of the multiple Mann-Whitney tests followed by Bonferroni adjustments between the HIV and healthy control group for each gated population is shown using P value significance codes found directly above: 0 *** 0.001 ** 0.01 * 0.05.

doi:10.1371/journal.pone.0137635.g003

Table 2. Non-parametric Spearman rank tests correlation analysis of the manual, sFLOCK and aFLOCK immunopathological populations to the clinical parameters.

	aFLOCK		sFLOCK		Manual	
	#	R (±SD) ^a	#	R (±SD) ^a	#	R (±SD) ^a
CD4 level	4	0.579 (0.131)	2	0.597 (0.191)	3	0.545 (0.037)
CD4%	5	0.622 (0.099)	2	0.633 (0.17)	3	0.586 (0.043)
CD8 level	1	0.384 (0.104)	1	0.385 (0.168)	2	0.404 (0.122)
CD8%	4	0.531 (0.09)	2	0.518 (0.149)	2	0.544 (0.119)
VL	1	0.276 (0.178)	0	0.296 (0.152)	1	0.304 (0.211)
CD4/CD8 ratio	5	0.634 (0.093)	3	0.64 (0.136)	3	0.617 (0.088)

^a Bonferroni corrections have been performed.

doi:10.1371/journal.pone.0137635.t002

CD38+, PD-1+, CD27+ and CD45RO+. At least one AIDS subject was an extreme outlier in all of these artificial generated clusters and these clusters were shown to be significantly more densely populated for HIV-infected subjects compared to the healthy controls (Fig 3). The majority of the clusters, identified by artificial FLOCK that correlated with the clinical parameters, were memory T cells with multi-dysfunctional traits. Cluster 15 (CD57loCD28hiCD38+PD-1+CD27+Tim-3loCD45ROhiHLA-DR+) was shown to best correlate with all clinical parameters in this data set, where it achieved a correlation of $Rho = -0.80$ (P value = 1.41×10^{-9}) with the CD4/CD8 ratio.

The T cell clusters 4, 10 and 12 significantly correlated with at least one of the clinical parameters for the sFLOCK dataset. These clusters of cells were CD28hi, PD-1+, CD27+, CD45ROhi and HLA-DR+, where AIDS outliers can be seen in Fig 3 for clusters 4 and 12. The activated/exhausted early-differentiated cluster 12 (CD57loCD28hiCD38loPD-1+CD27+Tim-3-CD45ROhiHLA-DR+) of sFLOCK correlated most significantly with the CD4/CD8 ratio ($Rho = -0.78$, P value = 5.75×10^{-9}). This shows that CD4+ T cells with a similar phenotype correlate with HIV disease progression independently of reference subjects. A noteworthy difference however, was that only two sFLOCK clusters correlated with the CD4% whereas three sFLOCK clusters correlated significantly with the CD4/CD8 ratio.

In summary, the manual analysis results showed that activation (CD4+CD38+HLA-DR+, CD4+HLA-DR+) and exhaustion (CD4+PD-1+) markers were primarily correlated with the laboratory parameters, whereas the results from the FLOCK analyses showed that it was specific activation and exhaustion profiles of early-differentiated memory clusters that were linked to the immunopathogenesis of HIV infection. In agreement to what we previously has observed [19], the CD4/CD8 ratio was shown to be the preeminent surrogate marker of pathological CD4+ T cells clusters identified with FLOCK. This data suggest that multidimensional clustering approaches like FLOCK is able to define and correlate specific CD4+ T cell clusters of relevance with HIV disease progression.

Discussion

Despite the great advances during the past 20 years in reducing AIDS-related mortality, standard therapies do not fully restore health or normal immune status in HIV-infected individuals. Despite many years on therapy, HIV-infected patients maintain elevated levels of immune activation and inflammation, which is linked to the increased incidence of cardiovascular diseases, bone disorders, cognitive impairment and other age-related diseases despite low viral loads and high CD4 counts [40]. Similarly, despite un-detectable viral load and relatively stable

high levels of CD4+ T cells, also elite controllers exhibit increased immune activation that is reversed by ART [41]. Thus, it still remains important to understand the dysfunctional changes of the immune system that persists in treated and un-treated HIV-infection.

Due to the myriad events that occur in HIV-infection, it remains highly important to understand the immunopathological changes at the single-cell level. FCM has long been the state-of-the-art method to study this phenomenon [42]. The expression of markers representing the activation (CD38, HLA-DR), exhaustion (PD-1, Tim-3), senescence (CD28, CD57) and memory differentiation (CD45RO, CD27) status of CD4+ T cells were here measured by FCM. Traditionally, FCM data has been analyzed using 1- or 2-dimensional plots sequentially. In this study, FLOCK [36], an algorithm that uses a density-based approach to identify biologically relevant clusters in the multidimensional space, was used to explore the immunopathogenesis of HIV. FLOCK was chosen for its ability to compare the occupation of populations between the two cohorts, the healthy controls and the HIV-infected individuals. During the FlowCAP challenge, FLOCK was shown to perform well in the completely automated challenge [30]. FLOCK was originally used to study B cell subsets found in human PBMC samples, where seventeen B cell subsets (including novel plasmablast subsets) were elucidated. FLOCK has since been used to study a range of different FCM datasets including an investigation into the response kinetics of T cells live-attenuated *S. Typhi* vaccine [36, 43–47]. To our knowledge, this is the first study that uses FLOCK to investigate a HIV dataset and provide this alternative approach to select a less biased reference for the algorithm.

A reference sample is a requirement for the FLOCK algorithm, from which the cell populations are determined and their centroids are applied to the remaining cohort samples to allow for cross-sample comparisons. A majority of the studies performed a FLOCK analysis of each of their samples and chose the best as a representative sample to define their centroids. The centroids have usually then been applied to all the subjects for a cross sample comparison [36, 45–47]. Although this method has been shown to be the typical way of analyzing FCM data using FLOCK, it imposes a level of subjectivity to the analysis, as one would have to determine the sample that one “feels” gives the best representation of the data. In this study, this normal approach of selecting the centroids was compared to a more automated method, where approximately 2% of each samples were concatenated to create an artificial reference. This artificial reference was used to select the centroids, after which the centroids were applied to all samples. The results showed that the automated method produced a larger number of clusters and a larger proportion of these clusters were shown to be immunopathological, *i.e.* different between the HIV-infected and healthy control subjects. This implies that the artificial reference provided a method of capturing a greater view of the pathogenesis distinguished in HIV-infected individuals. Henn et al. [43] used a similar method to prepare the centroids for cross-sample comparisons as the one presented in this article. They concatenated all of the samples in their study to define the centroids before applying them to all samples to study B cell responses in relation to vaccination to influenza. We chose only to use a proportion of the cells extracted from each sample to obtain a reference patient with approximately as many data points as the real patients. This artificial reference generated 25 independent clusters. When the resulting centroids were applied on all samples, twelve of the clusters were shown to be significantly different (after Bonferroni corrections) between the HIV-infected and healthy control subjects. Six of the immunopathological clusters were shown to correlate significantly with both the clinical parameters CD4% and CD4/CD8 ratio. In summary, the CD4/CD8 ratio was shown to correlate most significantly with the immunopathological clusters determined using the manual data analysis and both FLOCK results.

The results from the artificial reference (aFLOCK) were compared to the case of using a single representative individual as a reference for FLOCK (sFLOCK) as well as to the manual data

analysis. Interestingly, when all three methods were compared, aFLOCK appeared to improve the separation between healthy controls and HIV-infected individuals in comparison with sFLOCK or manual data analysis. In all three analyses, the first principle component separated the AIDS patients from the remaining HIV-infected individuals and healthy controls, whereas the second principle component was better at separating the HIV-infected individuals from the healthy controls. However, aFLOCK was shown to surpass at separating the distribution of HIV-infected individuals and healthy controls, when compared to the sFLOCK and manual data analysis.

In a previous study on the same cohort, a multi-parametric bioinformatics approach was utilized to determine the immunopathological CD4+ and CD8+ T cell subsets that correlated significantly with the routine laboratory parameters [19]. One of the methods used was Boolean gating, where an early-differentiated CD4+ T cell population expressing activation and exhaustion markers (CD57-CD28+CD38+PD-1+CD27+Tim-3-CD45RO+HLA-DR+) was best correlated with the CD4/CD8 ratio. Numerous previous observations implicate that the homeostatic failure of CD4+ T cell regeneration in chronic HIV infection, is a consequence of a highly dysfunctional and activated pool of early-differentiated CD4+ T cells [47]. This early-differentiated CD4+ T cell population is synonymous to the cluster that aFLOCK identified, (CD57loCD28hiCD38+PD-1+CD27+Tim-3loCD45ROhiHLA-DR+), that correlated most significantly with all the clinical parameters, particularly with the CD4/CD8 ratio. Cluster 12 (CD57loCD28hiCD38loPD-1+CD27+Tim-3-CD45ROhiHLA-DR+) of sFLOCK correlated most significantly to CD4/CD8 ratio with a correlation coefficient of -0.78. Interestingly, the sFLOCK, aFLOCK and Boolean populations that correlated most significantly with the clinical parameters were synonymous. The key difference is that the FLOCK results dispatch a more detailed phenotype of the populations and are not limited to the binary expression of a marker. This is important as it relays a more biologically correct representation of the immune system and offers a less time-consuming way without subjective gating of the populations.

A possible limitation of the study was the exclusion of healthy subjects from the artificial reference, where it would have been interesting to create the artificial reference based on the combined measurements of HIV-infected individuals and healthy controls together. However, we decided to concentrate on the HIV-relevant phenotypes in this study as these subjects most probably were going to generate more extreme populations than healthy controls. Notably some of the HIV-infected individuals were relatively healthy and therefore some of the less dysfunctional subsets should also be present. It is hard to predict how robust this method of 'selecting' a reference patient is when a much larger number of subjects are used to create the reference. However, the same issue of whether to pick a healthy, HIV+ or AIDS subject, as a single reference subject would have to be addressed when using the standard FLOCK analysis. It should be iterated that this method of selecting a reference sample could be generalized to analyses where cross-sample comparisons are to be made. A further limitation is the cross-sectional nature of this study, where multiple time points or estimated time from infection would have been highly desirable to *e.g.* determine how the CD4/CD8 ratio could predict disease progression or mortality in this cohort. Like previously discussed [19] and mentioned elsewhere [48], we here aimed to conduct an observational study on a highly ethnical diverse population of HIV-infected subjects to determine which traditional biomarker of HIV disease progression that was primarily associated with the dysfunctional CD4+ T cell clusters. Taken together, the measurement of one single time-point is of clinical interest, independently of assessment over time.

The CD4/CD8 ratio has typically been linked in elderly with a general state of immune dysfunction, where an inverted CD4/CD8 ratio is associated with short-time mortality [49–51]. Also in HIV infection the CD4/CD8 ratio has received renewed interest, particularly as it is

usually not normalized despite long-term therapy. Older studies have demonstrated that AIDS development in untreated HIV infection is highly predicted by longitudinal assessment of the CD4/CD8 ratio [48], whereas the CD4/CD8 ratio in treated individuals is related to the risk of the comorbidities [52]. In close linkage to our results, a recent report described that a low CD4/CD8 ratio during ART was associated with persistent immune activation and senescent profiles of T cells [53]. Whether the preserved immune activation profile might be a consequence or cause of a low CD4/CD8 ratio is hard to determine, but all these studies together suggest that the CD4/CD8 ratio could be used as an adequate biomarker for monitoring the state of immune dysfunction and morbidity/mortality in long-term treated HIV-infected subjects. In the wake of studies showing a clear benefit of ART to reduce mortality and transmission levels, possibly all affected individuals in developed countries will be treated in a near future. However, because of socio-economical aspects and other reasons, this seems to be a more distant prospect for developing countries, and therefore an informative biomarker of CD4+ T cell dysfunction and state of HIV disease could also serve as vital information in untreated subjects for a long time ahead. Other studies have demonstrated that the CD4 count is not a perfect predictor of HIV disease progression in HIV-infected individuals from developing countries [54, 55], further emphasizing the role of the CD4/CD8 ratio as a biomarker of interest also in untreated subjects in countries with the major burden of HIV infections.

In this study we have illustrated that the use of an automated clustering algorithm produced better results than the classical manual in separating the HIV-infected individuals and healthy controls. It is important to rely more on the automated methods, as they are less subjective than the manual gating methods. Supplemental to the FLOCK algorithm, we suggest an alternative method for selecting the reference subject for the cross-sample comparison to remove the subjectivity of selecting a reference subject. To sum up, with the growing number of parameters that can be measured with multiparametric flow cytometry it is important to use the automated methods to perform the data analysis as it has been shown in multiple studies that they are just as good, if not better than the classical data analyses.

Supporting Information

S1 Table. Description of the FLOCK population for the single HIV reference and artificial reference.

(PDF)

S2 Table. Clinical parameters correlating with the populations.

(PDF)

S3 Table. An in-depth overview of the multiple correlations analysis.

(XLSX)

Acknowledgments

We thank the study subjects for their participation as well as Marja Ahlqvist, Sofia Rydberg, and Elisabet Storgård for assistance with sample collection.

Author Contributions

Conceived and designed the experiments: JF MB OL ACK. Performed the experiments: JF MB KN. Analyzed the data: JF MB. Contributed reagents/materials/analysis tools: JF MB PN AS OL ACK. Wrote the paper: JF MB KN PN AS OL ACK.

References

1. Brenchley JM, Schacker TW, Ruff LE, Price DA, Taylor JH, Beilman GJ, et al. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J Exp Med*. 2004; 200(6):749–59. doi: [10.1084/jem.20040874](https://doi.org/10.1084/jem.20040874) PMID: [15365096](https://pubmed.ncbi.nlm.nih.gov/15365096/); PubMed Central PMCID: PMC2211962.
2. Mehandru S, Poles MA, Tenner-Racz K, Horowitz A, Hurley A, Hogan C, et al. Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. *J Exp Med*. 2004; 200(6):761–70. doi: [10.1084/jem.20041196](https://doi.org/10.1084/jem.20041196) PMID: [15365095](https://pubmed.ncbi.nlm.nih.gov/15365095/); PubMed Central PMCID: PMC2211967.
3. Deeks SG, Kitchen CM, Liu L, Guo H, Gascon R, Narvaez AB, et al. Immune activation set point during early HIV infection predicts subsequent CD4+ T-cell changes independent of viral load. *Blood*. 2004; 104(4):942–7. PMID: [15117761](https://pubmed.ncbi.nlm.nih.gov/15117761/).
4. Doitsh G, Galloway NL, Geng X, Yang Z, Monroe KM, Zepeda O, et al. Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. *Nature*. 2014; 505(7484):509–14. doi: [10.1038/nature12940](https://doi.org/10.1038/nature12940) PMID: [24356306](https://pubmed.ncbi.nlm.nih.gov/24356306/); PubMed Central PMCID: PMC4047036.
5. Cooper A, Garcia M, Petrovas C, Yamamoto T, Koup RA, Nabel GJ. HIV-1 causes CD4 cell death through DNA-dependent protein kinase during viral integration. *Nature*. 2013; 498(7454):376–9. doi: [10.1038/nature12274](https://doi.org/10.1038/nature12274) PMID: [23739328](https://pubmed.ncbi.nlm.nih.gov/23739328/).
6. Finkel TH, Tudor-Williams G, Banda NK, Cotton MF, Curiel T, Monks C, et al. Apoptosis occurs predominantly in bystander cells and not in productively infected cells of HIV- and SIV-infected lymph nodes. *Nat Med*. 1995; 1(2):129–34. PMID: [7585008](https://pubmed.ncbi.nlm.nih.gov/7585008/).
7. Muro-Cacho CA, Pantaleo G, Fauci AS. Analysis of apoptosis in lymph nodes of HIV-infected persons. Intensity of apoptosis correlates with the general state of activation of the lymphoid tissue and not with stage of disease or viral burden. *J Immunol*. 1995; 154(10):5555–66. PMID: [7730654](https://pubmed.ncbi.nlm.nih.gov/7730654/).
8. Kovacs JA, Lempicki RA, Sidorov IA, Adelsberger JW, Herpin B, Metcalf JA, et al. Identification of dynamically distinct subpopulations of T lymphocytes that are differentially affected by HIV. *J Exp Med*. 2001; 194(12):1731–41. PMID: [11748275](https://pubmed.ncbi.nlm.nih.gov/11748275/); PubMed Central PMCID: PMC2193579.
9. McCune JM, Hanley MB, Cesar D, Halvorsen R, Hoh R, Schmidt D, et al. Factors influencing T-cell turnover in HIV-1-seropositive patients. *The Journal of clinical investigation*. 2000; 105(5):R1–8. PMID: [10712441](https://pubmed.ncbi.nlm.nih.gov/10712441/).
10. Mohri H, Perelson AS, Tung K, Ribeiro RM, Ramratnam B, Markowitz M, et al. Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *J Exp Med*. 2001; 194(9):1277–87. PMID: [11696593](https://pubmed.ncbi.nlm.nih.gov/11696593/); PubMed Central PMCID: PMC2195973.
11. Giorgi JV, Detels R. T-cell subset alterations in HIV-infected homosexual men: NIAID Multicenter AIDS cohort study. *Clinical immunology and immunopathology*. 1989; 52(1):10–8. PMID: [2656013](https://pubmed.ncbi.nlm.nih.gov/2656013/).
12. Giorgi JV, Lyles RH, Matud JL, Yamashita TE, Mellors JW, Hultin LE, et al. Predictive value of immunologic and virologic markers after long or short duration of HIV-1 infection. *J Acquir Immune Defic Syndr*. 2002; 29(4):346–55. PMID: [11917238](https://pubmed.ncbi.nlm.nih.gov/11917238/).
13. Hazenberg MD, Stuart JW, Otto SA, Borleffs JC, Boucher CA, de Boer RJ, et al. T-cell division in human immunodeficiency virus (HIV)-1 infection is mainly due to immune activation: a longitudinal analysis in patients before and during highly active antiretroviral therapy (HAART). *Blood*. 2000; 95(1):249–55. PMID: [10607709](https://pubmed.ncbi.nlm.nih.gov/10607709/).
14. Sousa AE, Carneiro J, Meier-Schellersheim M, Grossman Z, Victorino RM. CD4 T cell depletion is linked directly to immune activation in the pathogenesis of HIV-1 and HIV-2 but only indirectly to the viral load. *J Immunol*. 2002; 169(6):3400–6. Epub 2002/09/10. PMID: [12218162](https://pubmed.ncbi.nlm.nih.gov/12218162/).
15. Hunt PW, Cao HL, Muzoora C, Ssewanyana I, Bennett J, Emenyonu N, et al. Impact of CD8+ T-cell activation on CD4+ T-cell recovery and mortality in HIV-infected Ugandans initiating antiretroviral therapy. *AIDS*. 2011; 25(17):2123–31. doi: [10.1097/QAD.0b013e32834c4ac1](https://doi.org/10.1097/QAD.0b013e32834c4ac1) PMID: [21881481](https://pubmed.ncbi.nlm.nih.gov/21881481/); PubMed Central PMCID: PMC3480326.
16. Day CL, Kaufmann DE, Kiepiela P, Brown JA, Moodley ES, Reddy S, et al. PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. *Nature*. 2006; 443(7109):350–4. Epub 2006/08/22. doi: [10.1038/nature05115](https://doi.org/10.1038/nature05115) PMID: [16921384](https://pubmed.ncbi.nlm.nih.gov/16921384/).
17. Jones RB, Ndhlovu LC, Barbour JD, Sheth PM, Jha AR, Long BR, et al. Tim-3 expression defines a novel population of dysfunctional T cells with highly elevated frequencies in progressive HIV-1 infection. *J Exp Med*. 2008; 205(12):2763–79. doi: [10.1084/jem.20081398](https://doi.org/10.1084/jem.20081398) PMID: [19001139](https://pubmed.ncbi.nlm.nih.gov/19001139/); PubMed Central PMCID: PMC2585847.
18. Trautmann L, Janbazian L, Chomont N, Said EA, Gimmig S, Bessette B, et al. Upregulation of PD-1 expression on HIV-specific CD8+ T cells leads to reversible immune dysfunction. *Nat Med*. 2006; 12(10):1198–202. Epub 2006/08/19. doi: [10.1038/nm1482](https://doi.org/10.1038/nm1482) PMID: [16917489](https://pubmed.ncbi.nlm.nih.gov/16917489/).

19. Buggert M, Frederiksen J, Noyan K, Svard J, Barqasho B, Sonnerborg A, et al. Multiparametric bioinformatics distinguish the CD4/CD8 ratio as a suitable laboratory predictor of combined T cell pathogenesis in HIV infection. *J Immunol*. 2014; 192(5):2099–108. Epub 2014/02/05. doi: [10.4049/jimmunol.1302596](https://doi.org/10.4049/jimmunol.1302596) PMID: [24493822](https://pubmed.ncbi.nlm.nih.gov/24493822/).
20. Eller MA, Blom KG, Gonzalez VD, Eller LA, Naluyima P, Laeyendecker O, et al. Innate and adaptive immune responses both contribute to pathological CD4 T cell activation in HIV-1 infected Ugandans. *PLoS One*. 2011; 6(4):e18779. doi: [10.1371/journal.pone.0018779](https://doi.org/10.1371/journal.pone.0018779) PMID: [21526194](https://pubmed.ncbi.nlm.nih.gov/21526194/); PubMed Central PMCID: PMC3079731.
21. Desai S, Landay A. Early immune senescence in HIV disease. *Current HIV/AIDS reports*. 2010; 7(1):4–10. doi: [10.1007/s11904-009-0038-4](https://doi.org/10.1007/s11904-009-0038-4) PMID: [20425052](https://pubmed.ncbi.nlm.nih.gov/20425052/); PubMed Central PMCID: PMC3739442.
22. Aghaepour N, Chattopadhyay PK, Ganesan A, O'Neill K, Zare H, Jalali A, et al. Early immunologic correlates of HIV protection can be identified from computational analysis of complex multivariate T-cell flow cytometry assays. *Bioinformatics*. 2012; 28(7):1009–16. doi: [10.1093/bioinformatics/bts082](https://doi.org/10.1093/bioinformatics/bts082) PMID: [22383736](https://pubmed.ncbi.nlm.nih.gov/22383736/); PubMed Central PMCID: PMC3315712.
23. Freer G, Rindi L. Intracellular cytokine detection by fluorescence-activated flow cytometry: basic principles and recent advances. *Methods*. 2013; 61(1):30–8. doi: [10.1016/j.ymeth.2013.03.035](https://doi.org/10.1016/j.ymeth.2013.03.035) PMID: [23583887](https://pubmed.ncbi.nlm.nih.gov/23583887/).
24. O'Donnell EA, Ernst DN, Hingorani R. Multiparameter flow cytometry: advances in high resolution analysis. *Immune network*. 2013; 13(2):43–54. doi: [10.4110/in.2013.13.2.43](https://doi.org/10.4110/in.2013.13.2.43) PMID: [23700394](https://pubmed.ncbi.nlm.nih.gov/23700394/); PubMed Central PMCID: PMC3659255.
25. Pedreira CE, Costa ES, Lecrevisse Q, van Dongen JJ, Orfao A, EuroFlow C. Overview of clinical flow cytometry data analysis: recent advances and future challenges. *Trends in biotechnology*. 2013; 31(7):415–25. doi: [10.1016/j.tibtech.2013.04.008](https://doi.org/10.1016/j.tibtech.2013.04.008) PMID: [23746659](https://pubmed.ncbi.nlm.nih.gov/23746659/).
26. Chattopadhyay PK, Roederer M. Good cell, bad cell: flow cytometry reveals T-cell subsets important in HIV disease. *Cytometry A*. 2010; 77(7):614–22. doi: [10.1002/cyto.a.20905](https://doi.org/10.1002/cyto.a.20905) PMID: [20583275](https://pubmed.ncbi.nlm.nih.gov/20583275/); PubMed Central PMCID: PMC2907059.
27. Qiu P, Simonds EF, Bendall SC, Gibbs KD Jr., Bruggner RV, Linderman MD, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nature biotechnology*. 2011; 29(10):886–91. doi: [10.1038/nbt.1991](https://doi.org/10.1038/nbt.1991) PMID: [21964415](https://pubmed.ncbi.nlm.nih.gov/21964415/); PubMed Central PMCID: PMC3196363.
28. Amadori A, Zamarchi R, Chieco-Bianchi L. CD4: CD8 ratio and HIV infection: the "tap-and-drain" hypothesis. *Immunology today*. 1996; 17(9):414–7. PMID: [8854558](https://pubmed.ncbi.nlm.nih.gov/8854558/).
29. Chattopadhyay PK, Roederer M. Cytometry: today's technology and tomorrow's horizons. *Methods*. 2012; 57(3):251–8. doi: [10.1016/j.ymeth.2012.02.009](https://doi.org/10.1016/j.ymeth.2012.02.009) PMID: [22391486](https://pubmed.ncbi.nlm.nih.gov/22391486/); PubMed Central PMCID: PMC3374038.
30. Aghaepour N, Finak G, Flow CAPC, Consortium D, Hoos H, Mosmann TR, et al. Critical assessment of automated flow cytometry data analysis techniques. *Nature methods*. 2013; 10(3):228–38. doi: [10.1038/nmeth.2365](https://doi.org/10.1038/nmeth.2365) PMID: [23396282](https://pubmed.ncbi.nlm.nih.gov/23396282/); PubMed Central PMCID: PMC3906045.
31. O'Neill K, Aghaepour N, Spidlen J, Brinkman R. Flow cytometry bioinformatics. *PLoS computational biology*. 2013; 9(12):e1003365. doi: [10.1371/journal.pcbi.1003365](https://doi.org/10.1371/journal.pcbi.1003365) PMID: [24363631](https://pubmed.ncbi.nlm.nih.gov/24363631/); PubMed Central PMCID: PMC3867282.
32. Baumgarth N, Roederer M. A practical approach to multicolor flow cytometry for immunophenotyping. *J Immunol Methods*. 2000; 243(1–2):77–97. PMID: [10986408](https://pubmed.ncbi.nlm.nih.gov/10986408/).
33. Lugli E, Roederer M, Cossarizza A. Data analysis in flow cytometry: the future just started. *Cytometry A*. 2010; 77(7):705–13. doi: [10.1002/cyto.a.20901](https://doi.org/10.1002/cyto.a.20901) PMID: [20583274](https://pubmed.ncbi.nlm.nih.gov/20583274/); PubMed Central PMCID: PMC2909632.
34. Zare H, Shoostari P, Gupta A, Brinkman RR. Data reduction for spectral clustering to analyze high throughput flow cytometry data. *BMC Bioinformatics*. 2010; 11:403. doi: [10.1186/1471-2105-11-403](https://doi.org/10.1186/1471-2105-11-403) PMID: [20667133](https://pubmed.ncbi.nlm.nih.gov/20667133/); PubMed Central PMCID: PMC2923634.
35. Aghaepour N, Nikolic R, Hoos HH, Brinkman RR. Rapid cell population identification in flow cytometry data. *Cytometry A*. 2011; 79(1):6–13. doi: [10.1002/cyto.a.21007](https://doi.org/10.1002/cyto.a.21007) PMID: [21182178](https://pubmed.ncbi.nlm.nih.gov/21182178/); PubMed Central PMCID: PMC3137288.
36. Qian Y, Wei C, Eun-Hyung Lee F, Campbell J, Halliley J, Lee JA, et al. Elucidation of seventeen human peripheral blood B-cell subsets and quantification of the tetanus response using a density-based method for the automated identification of cell populations in multidimensional flow cytometry data. *Cytometry Part B, Clinical cytometry*. 2010; 78 Suppl 1:S69–82. doi: [10.1002/cyto.b.20554](https://doi.org/10.1002/cyto.b.20554) PMID: [20839340](https://pubmed.ncbi.nlm.nih.gov/20839340/); PubMed Central PMCID: PMC3084630.
37. Buggert M, Tauriainen J, Yamamoto T, Frederiksen J, Ivarsson MA, Michaelsson J, et al. T-bet and Eomes are differentially linked to the exhausted phenotype of CD8+ T cells in HIV infection. *PLoS pathogens*. 2014; 10(7):e1004251. Epub 2014/07/18. doi: [10.1371/journal.ppat.1004251](https://doi.org/10.1371/journal.ppat.1004251) PMID: [25032686](https://pubmed.ncbi.nlm.nih.gov/25032686/); PubMed Central PMCID: PMC4102564.

38. Norstrom MM, Buggert M, Tauriainen J, Hartogensis W, Prosperi MC, Wallet MA, et al. Combination of immune and viral factors distinguishes low-risk versus high-risk HIV-1 disease progression in HLA-B*5701 subjects. *J Virol*. 2012; 86(18):9802–16. Epub 2012/07/05. doi: [10.1128/JVI.01165-12](https://doi.org/10.1128/JVI.01165-12) PMID: [22761389](https://pubmed.ncbi.nlm.nih.gov/22761389/); PubMed Central PMCID: PMC3446568.
39. R RDCT. R: A Language and Environment for Statistical Computing. R Found Stat Comput 1: 409. Vienna, Austria 2011.
40. Deeks SG. HIV infection, inflammation, immunosenescence, and aging. *Annual review of medicine*. 2011; 62:141–55. Epub 2010/11/26. doi: [10.1146/annurev-med-042909-093756](https://doi.org/10.1146/annurev-med-042909-093756) PMID: [21090961](https://pubmed.ncbi.nlm.nih.gov/21090961/).
41. Hatano H, Yukl SA, Ferre AL, Graf EH, Somsouk M, Sinclair E, et al. Prospective antiretroviral treatment of asymptomatic, HIV-1 infected controllers. *PLoS pathogens*. 2013; 9(10):e1003691. doi: [10.1371/journal.ppat.1003691](https://doi.org/10.1371/journal.ppat.1003691) PMID: [24130489](https://pubmed.ncbi.nlm.nih.gov/24130489/); PubMed Central PMCID: PMC3795031.
42. Cossarizza A, De Biasi S, Gibellini L, Bianchini E, Bartolomeo R, Nasi M, et al. Cytometry, immunology, and HIV infection: three decades of strong interactions. *Cytometry A*. 2013; 83(8):680–91. doi: [10.1002/cyto.a.22318](https://doi.org/10.1002/cyto.a.22318) PMID: [23788450](https://pubmed.ncbi.nlm.nih.gov/23788450/).
43. Henn AD, Wu S, Qiu X, Ruda M, Stover M, Yang H, et al. High-resolution temporal response patterns to influenza vaccine reveal a distinct human plasma cell gene signature. *Scientific reports*. 2013; 3:2327. doi: [10.1038/srep02327](https://doi.org/10.1038/srep02327) PMID: [23900141](https://pubmed.ncbi.nlm.nih.gov/23900141/); PubMed Central PMCID: PMC3728595.
44. Kaminski DA, Wei C, Qian Y, Rosenberg AF, Sanz I. Advances in human B cell phenotypic profiling. *Frontiers in immunology*. 2012; 3:302. doi: [10.3389/fimmu.2012.00302](https://doi.org/10.3389/fimmu.2012.00302) PMID: [23087687](https://pubmed.ncbi.nlm.nih.gov/23087687/); PubMed Central PMCID: PMC3467643.
45. Lelic A, Verschoor CP, Ventresca M, Parsons R, Eveleigh C, Bowdish D, et al. The polyfunctionality of human memory CD8+ T cells elicited by acute and chronic virus infections is not influenced by age. *PLoS pathogens*. 2012; 8(12):e1003076. doi: [10.1371/journal.ppat.1003076](https://doi.org/10.1371/journal.ppat.1003076) PMID: [23271970](https://pubmed.ncbi.nlm.nih.gov/23271970/); PubMed Central PMCID: PMC3521721.
46. McArthur MA, Sztejn MB. Heterogeneity of multifunctional IL-17A producing S. Typhi-specific CD8+ T cells in volunteers following Ty21a typhoid immunization. *PLoS One*. 2012; 7(6):e38408. doi: [10.1371/journal.pone.0038408](https://doi.org/10.1371/journal.pone.0038408) PMID: [22679502](https://pubmed.ncbi.nlm.nih.gov/22679502/); PubMed Central PMCID: PMC3367967.
47. McArthur MA, Sztejn MB. Unexpected heterogeneity of multifunctional T cells in response to superantigen stimulation in humans. *Clinical immunology*. 2013; 146(2):140–52. doi: [10.1016/j.clim.2012.12.003](https://doi.org/10.1016/j.clim.2012.12.003) PMID: [23333555](https://pubmed.ncbi.nlm.nih.gov/23333555/); PubMed Central PMCID: PMC3565224.
48. Taylor JM, Fahey JL, Detels R, Giorgi JV. CD4 percentage, CD4 number, and CD4:CD8 ratio in HIV infection: which to choose and how to use. *J Acquir Immune Defic Syndr*. 1989; 2(2):114–24. PMID: [2495346](https://pubmed.ncbi.nlm.nih.gov/2495346/).
49. Strindhall J, Nilsson BO, Lofgren S, Ernerudh J, Pawelec G, Johansson B, et al. No Immune Risk Profile among individuals who reach 100 years of age: findings from the Swedish NONA immune longitudinal study. *Experimental gerontology*. 2007; 42(8):753–61. doi: [10.1016/j.exger.2007.05.001](https://doi.org/10.1016/j.exger.2007.05.001) PMID: [17606347](https://pubmed.ncbi.nlm.nih.gov/17606347/).
50. Wikby A, Johansson B, Ferguson F, Olsson J. Age-related changes in immune parameters in a very old population of Swedish people: a longitudinal study. *Experimental gerontology*. 1994; 29(5):531–41. PMID: [7828662](https://pubmed.ncbi.nlm.nih.gov/7828662/).
51. Wikby A, Maxson P, Olsson J, Johansson B, Ferguson FG. Changes in CD8 and CD4 lymphocyte subsets, T cell proliferation responses and non-survival in the very old: the Swedish longitudinal OCTO-immune study. *Mechanisms of ageing and development*. 1998; 102(2–3):187–98. PMID: [9720651](https://pubmed.ncbi.nlm.nih.gov/9720651/).
52. Kuller LH, Tracy R, Bellosso W, De Wit S, Drummond F, Lane HC, et al. Inflammatory and coagulation biomarkers and mortality in patients with HIV infection. *PLoS medicine*. 2008; 5(10):e203. doi: [10.1371/journal.pmed.0050203](https://doi.org/10.1371/journal.pmed.0050203) PMID: [18942885](https://pubmed.ncbi.nlm.nih.gov/18942885/); PubMed Central PMCID: PMC2570418.
53. Serrano-Villar S, Sainz T, Lee SA, Hunt PW, Sinclair E, Shacklett BL, et al. HIV-infected individuals with low CD4/CD8 ratio despite effective antiretroviral therapy exhibit altered T cell subsets, heightened CD8+ T cell activation, and increased risk of non-AIDS morbidity and mortality. *PLoS pathogens*. 2014; 10(5):e1004078. doi: [10.1371/journal.ppat.1004078](https://doi.org/10.1371/journal.ppat.1004078) PMID: [24831517](https://pubmed.ncbi.nlm.nih.gov/24831517/); PubMed Central PMCID: PMC4022662.
54. Anglaret X, Diagbouga S, Mortier E, Meda N, Verge-Valette V, Sylla-Koko F, et al. CD4+ T-lymphocyte counts in HIV infection: are European standards applicable to African patients? *Journal of acquired immune deficiency syndromes and human retrovirology: official publication of the International Retrovirology Association*. 1997; 14(4):361–7. PMID: [9111479](https://pubmed.ncbi.nlm.nih.gov/9111479/).
55. Anglaret X, Sylla-Koko F, Diagbouga S, Combe P, Van De Perre P, Dabis F. CD4 count and CD4% in African HIV-infected people. *International journal of epidemiology*. 1998; 27(5):928–9. PMID: [9839756](https://pubmed.ncbi.nlm.nih.gov/9839756/).